

Uniform Generation in Trace Monoids*

Samy Abbes[†]

Jean Mairesse[‡]

2014

Abstract

We consider the problem of random uniform generation of traces (the elements of a free partially commutative monoid) in light of the uniform measure on the boundary at infinity of the associated monoid. We obtain a product decomposition of the uniform measure at infinity if the trace monoid has several irreducible components—a case where other notions such as Parry measures, are not defined. Random generation algorithms are then examined.

Keywords: trace monoid, uniform generation, Möbius polynomial

1—Introduction

Uniform generation of finite-size combinatorial objects consists in the design of a randomized algorithm that takes an integer k as input, and returns an object of size k , such that each object of size k has equal probability to be produced. This problem has been considered for many classes of objects from computer science or discrete mathematics: words, trees, graphs are examples. Several general approaches exist: recursive methods [11], the Markov chain Monte-Carlo method with coupling from the past [12], or the Boltzmann sampler [10]. Other recent approaches share a common guideline, namely first considering a notion of *uniform measure on infinite objects* in order to gain, afterwards, information on the uniform distributions on finite objects. The theory of random planar graphs is an example of application of this idea. In this paper, we investigate the uniform generation of traces (elements of a trace monoid) and we base our approach on the notion of *uniform measure on infinite traces*.

Given an independence pair (A, I) , where I is an irreflexive and symmetric relation on the finite alphabet A , the associated trace monoid $\mathcal{M} = \mathcal{M}(A, I)$ contains all congruence classes of the free monoid A^* , modulo equivalences of the form $ab = ba$ for all $(a, b) \in I$, see [6, 8]. Elements of \mathcal{M} are called *traces*. Trace monoids are ubiquitous in Combinatorics, see [18]. They are also one of

*Published in the proceedings of *Mathematical Foundations of Computer Science 2015, Milano*

[†]Université Paris Diderot/PPS CNRS UMR 7126 and IRISA/INRIA CNRS UMR 6074
samy.abbes@univ-paris-diderot.fr

[‡]UPMC/LIP6 CNRS UMR 7606 jean.mairesse@lip6.fr

the most basic models of concurrency under a partial order semantics [9]. Uniform generation of traces is thus a fundamental question with possible applications in probabilistic model checking of concurrent systems. Since our concern is with partial order semantics, it differs from the sequential approach which targets uniform generation of linear executions in models of concurrency [5].

Consider a trace monoid \mathcal{M} , and, for each integer $k \geq 0$, the finite set $\mathcal{M}_k = \{x \in \mathcal{M} : |x| = k\}$. Let $\nu_{\mathcal{M}_k}$ be the uniform distribution over \mathcal{M}_k . A crucial observation is that the probability measures $(\nu_{\mathcal{M}_k})_{k \in \mathbb{N}}$ are *not* consistent. Consequently, the uniform measures $\nu_{\mathcal{M}_k}$ cannot be reached by a recursive sampling of the form $x_1 \cdot \dots \cdot x_k \in \mathcal{M}_k$, with the x_i 's being sampled independently and according to some common distribution over A .

To overcome the difficulty, several steps are necessary. First, we consider the *uniform measure at infinity* for \mathcal{M} , a notion introduced in [2] for irreducible trace monoids, and extended here to the general case. Second, we prove a realization result for the uniform measure at infinity by means of a Markov chain on a combinatorial sub-shift. Last, we apply the results to the uniform sampling of finite traces. None of the three steps is straightforward. Besides standard uniform sampling, it turns out that evaluating the uniform average cost or reward associated with traces can be done in an efficient way.

An original feature of our approach is to define the measure at infinity for general trace monoids and not only for irreducible ones. We show that the uniform measure at infinity of a *reducible* trace monoid decomposes as a product of measures on irreducible components—contrasting with uniform distribution at finite horizon. In general, the uniform measure at infinity charges the *infinite* traces of the “largest” components of the monoid, and charges the *finite* traces of the “smallest” components.

Another, different but related, notion of ‘uniform measure’ exists: the Parry measure which is a uniform measure on bi-infinite sequences of an *irreducible* sofic sub-shift [16, 13]. The construction can be applied to trace monoids, defining a ‘uniform measure’ on bi-infinite traces, but only for irreducible trace monoids. Here we focus on single sided infinite traces instead of bi-infinite ones, and this approach allows to relax the irreducibility assumption, and to construct a uniform measure at infinity for a general trace monoid. In case the trace monoid is irreducible, we provide a precise comparison between the Parry measure, restricted to single sided infinite traces, and our uniform measure at infinity. The latter turns out to be a non-stationary version of the former. Another important point is that our approach reveals the combinatorial structure hidden in the uniform measure at infinity (and in the Parry measure).

The outline of the paper is the following. We first focus in a warm-up section (§ 2) on the case of two commuting alphabets. Relaxing the commutativity assumption, we arrive to trace monoids in § 3. The purpose of § 4 is twofold: first, to compare the uniform measure with the Parry measure; and second, to examine applications to the uniform sampling of finite traces.

2—Warm-up: uniform measure for commuting alphabets

Let A and B be two alphabets and let \mathcal{M} be the product monoid $\mathcal{M} = A^* \times B^*$. The size of $u = (x, y)$ in \mathcal{M} is $|u| = |x| + |y|$. Let $\partial A^* = A^{\mathbb{N}}$ be the set of

infinite A -words, let $\overline{A^*} = A^* \cup A^{\mathbb{N}}$, and similarly for ∂B^* and $\overline{B^*}$. Define:

$$\partial \mathcal{M} = \{(\xi, \zeta) \in \overline{A^*} \times \overline{B^*} : |\xi| + |\zeta| = \infty\}, \quad \overline{\mathcal{M}} = \mathcal{M} \cup \partial \mathcal{M}.$$

Clearly one has $\partial \mathcal{M} = (\overline{A^*} \times \overline{B^*}) - (A^* \times B^*)$ and $\overline{\mathcal{M}} = \overline{A^*} \times \overline{B^*}$. Both $\overline{A^*}$ and $\overline{B^*}$ are equipped with the natural prefix orderings, and $\overline{\mathcal{M}}$ is equipped with the product ordering, denoted by \leq . For $u \in \mathcal{M}$, we put:

$$\uparrow u = \{v \in \overline{\mathcal{M}} : u \leq v\}, \quad \uparrow u = \{\xi \in \partial \mathcal{M} : u \leq \xi\}.$$

Let $p_0 = 1/|A|$ and $q_0 = 1/|B|$. Without loss of generality, we assume that $|A| \geq |B|$, hence $p_0 \leq q_0$.

• **Lemma 1**—*For each real number $p \in (0, p_0]$, there exists a unique probability measure ν_p on $\overline{A^*}$ such that $\nu_p(\uparrow x) = p^{|x|}$ holds for all $x \in A^*$. We have:*

$$\forall p \in (0, p_0) \quad \nu_p(A^*) = 1, \quad \nu_{p_0}(\partial A^*) = 1.$$

The probability measures ν_p in Lemma 1 are called *sub-uniform* measures of parameter p over $\overline{A^*}$. The measure ν_{p_0} is the classical uniform measure on ∂A^* which satisfies $\nu_{p_0}(\uparrow x) = p_0^{|x|}$ for all $x \in A^*$.

For each integer $k \geq 0$, let $\nu_{\mathcal{M}_k}$ denote the uniform distribution on $\mathcal{M}_k = \{(x, y) \in \mathcal{M} : |(x, y)| = k\}$. Since $|A| \geq |B|$, an element $(x, y) \in \mathcal{M}_k$ sampled according to $\nu_{\mathcal{M}_k}$ is more likely to satisfy $|x| \geq |y|$ than the opposite. In the limit, it is natural to expect that infinite elements on the B side are not charged at all, except if $|A| = |B|$. This is made precise in the following result.

• **Theorem 1**—*Let ν_A and ν_B be the sub-uniform measures of parameter $p_0 = 1/|A|$ over $\overline{A^*}$ and $\overline{B^*}$ respectively. The sequence $(\nu_{\mathcal{M}_k})_{k \geq 0}$ converges weakly to the product measure $\nu = \nu_A \otimes \nu_B$.*

We have: $\nu(\uparrow(x, y)) = p_0^{|x|+|y|}$ for all $(x, y) \in \mathcal{M}$; and $\nu(\partial A^ \times B^*) = 1$ if $|A| > |B|$, whereas $\nu(\partial A^* \times \partial B^*) = 1$ if $|A| = |B|$.*

We say that the measure ν described in Th. 1 is the *uniform* measure on $\partial \mathcal{M}$. We have the following “realization” result for ν .

• **Theorem 2**—*Let $(a_n)_{n \in \mathbb{N}}$ be a sequence of i.i.d. and uniform random variables (r.v.) over A . Let b_0 be a r.v. over $B \cup \{1_{B^*}\}$, where 1_{B^*} is the identity element of B^* , and with the following law:*

$$\forall b \in B \quad \mathbb{P}(b_0 = b) = p_0 = 1/|A|, \quad \mathbb{P}(b_0 = 1_{B^*}) = 1 - p_0/q_0 = 1 - |B|/|A|.$$

Consider $(b_n)_{n \in \mathbb{N}}$ sampled independently in $B \cup \{1_{B^}\}$, each b_n with the same law as b_0 , but only until it reaches 1_{B^*} , after which b_n is constant equal to 1_{B^*} . Finally, set $u_k \in \mathcal{M}$ for all integers $k \geq 0$ by:*

$$x_k = a_0 \cdot \dots \cdot a_{k-1} \in A^*, \quad y_k = b_0 \cdot \dots \cdot b_{k-1} \in B^*, \quad u_k = (x_k, y_k) \in \mathcal{M}.$$

Then $(u_k)_{k \in \mathbb{N}}$ converges in law towards ν . Furthermore, the random variable $\bigvee_{k \geq 0} u_k \in \partial \mathcal{M}$ is distributed according to ν .

Observe that 1_{B^*} will eventually appear in the sequence $(b_n)_{n \in \mathbb{N}}$ with probability 1 if and only if $p_0 < q_0$. In this case, $(y_n)_{n \in \mathbb{N}}$ is eventually equal to a constant element of B^* with probability 1. This is consistent with Theorem 1. Observe also that $(a_n, b_n)_{n \in \mathbb{N}}$ forms a product Markov chain on $A \times (B \cup \{1_{B^*}\})$.

Both results stated in Ths. 1 and 2 are particular cases of corresponding results for trace monoids, as we will see next.

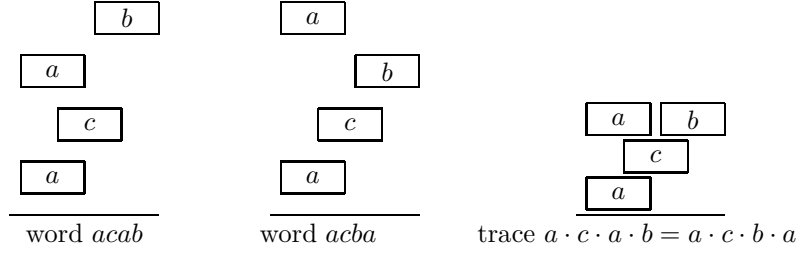


Figure 1: Two congruent words and the resulting heap (trace)

3—Uniform and sub-uniform measures for trace monoids

Basics on trace monoids. Let A be a finite alphabet equipped with an irreflexive and symmetric relation $I \subseteq A \times A$, called an *independence relation*. The pair (A, I) is called an *independence pair*. Let \mathcal{I} be the congruence relation on the free monoid A^* generated by the collection of pairs (ab, ba) for (a, b) ranging over I . The *trace monoid* $\mathcal{M} = \mathcal{M}(A, I)$ is defined as the quotient monoid $\mathcal{M} = A^*/\mathcal{I}$, see [6, 18, 8]. The elements of \mathcal{M} are called *traces*. The identity element in the monoid is called the *empty trace*, denoted “ $1_{\mathcal{M}}$ ”, and the concatenation is denoted with the dot “ \cdot ”.

The length of a trace u is well defined as the length of any of its representative words and is denoted by $|u|$. The left divisibility relation on \mathcal{M} is a partial order, denoted by “ \leq ” and defined by: $u \leq v \iff \exists w \quad v = u \cdot w$.

An intuitive representation of traces is given by Viennot’s *heap of pieces* interpretation of a trace monoid [18]. We illustrate in Fig. 1 the heap of pieces interpretation for the monoid $\mathcal{M}(A, I)$ with $A = \{a, b, c\}$ and $I = \{(a, b), (b, a)\}$.

The length of traces corresponds to the number of pieces in a heap. The relation $u \leq v$ corresponds to u being seen at bottom as a sub-heap of heap v .

The product monoid $A^* \times B^*$ from § 2 is isomorphic to the trace monoid $\mathcal{M}(\Sigma, I)$, where $\Sigma = A \cup B$ with A and B being considered as disjoint, and $I = (A \times B) \cup (B \times A)$.

Cliques and height of traces. Recall that a *clique* of a graph is a complete subgraph (by convention, the empty graph is a clique). We may view (A, I) as a graph. Given a clique c of (A, I) , the product $a_1 \cdot \dots \cdot a_j \in \mathcal{M}$ is independent of the enumeration (a_1, \dots, a_j) of the vertices composing c . We say that $a_1 \cdot \dots \cdot a_j$ is a *clique* of \mathcal{M} . Let \mathcal{C} denote the set of cliques, including the empty clique $1_{\mathcal{M}}$. As heaps of pieces, cliques correspond to flat heaps, or horizontal layers.

Traces are known to admit a canonical normal form, defined as follows [6]. Say that two non-empty cliques c, c' are *Cartier-Foata admissible*, denoted by $c \rightarrow c'$, whenever they satisfy: $\forall a \in c' \quad \exists b \in c \quad (b, a) \notin I$. For every non empty trace $u \in \mathcal{M}$, there exists a unique integer $n > 0$ and a unique sequence (c_1, \dots, c_n) of non-empty cliques such that: (1) $u = c_1 \cdot \dots \cdot c_n$; and (2) $c_i \rightarrow c_{i+1}$ holds for all $i \in \{1, \dots, n-1\}$. The integer n is called the *height* of u , denoted by $n = \tau(u)$. By convention, we put $\tau(1_{\mathcal{M}}) = 0$. The sequence (c_1, \dots, c_n) is called the *Cartier-Foata normal form* or *decomposition* of u . In the heap interpretation, the normal form corresponds to the sequence

of horizontal layers that compose a heap u , and the height $\tau(u)$ corresponds to the number of horizontal layers.

A useful device is the notion of *topping* of traces, defined as follows: for each integer $n \geq 0$, the n -*topping* is the mapping $\kappa_n : \mathcal{M} \rightarrow \mathcal{M}$ defined by $\kappa_n(u) = c_1 \cdot \dots \cdot c_n$, where $c_1 \rightarrow \dots \rightarrow c_p$ is the Cartier-Foata decomposition of u , and where $c_i = 1_{\mathcal{M}}$ if $i > p$.

Boundary. Elementary cylinders. Let $\mathfrak{C} = \mathcal{C} \setminus \{1_{\mathcal{M}}\}$ denote the set of non-empty cliques. Traces of \mathcal{M} are in bijection with finite paths of the automaton $(\mathfrak{C}, \rightarrow)$, where all states are both initial and final. Denote by $\partial\mathcal{M}$ the set of infinite paths in the automaton $(\mathfrak{C}, \rightarrow)$. We call $\partial\mathcal{M}$ the *boundary at infinity*, or simply the *boundary*, of monoid \mathcal{M} , and we put $\overline{\mathcal{M}} = \mathcal{M} \cup \partial\mathcal{M}$. Elements of $\partial\mathcal{M}$ are called *infinite traces*, and, by contrast, elements of \mathcal{M} might be called *finite traces*.

By construction, an infinite trace is given as an infinite sequence $\xi = (c_1, c_2, \dots)$ of non-empty cliques such that $c_i \rightarrow c_{i+1}$ holds for all integers $i \geq 1$. Note that the topping operations extend naturally to $\kappa_n : \overline{\mathcal{M}} \rightarrow \mathcal{M}$, defined by $\kappa_n(\xi) = c_1 \cdot \dots \cdot c_n$, for $\xi = (c_1, c_2, \dots)$.

We wish to extend the partial order relation \leq from \mathcal{M} to $\overline{\mathcal{M}}$. For this, we first recall the following result [2, Cor. 4.2]: for $u, v \in \mathcal{M}$, if $n = \tau(u)$, then $u \leq v \iff u \leq \kappa_n(v)$. Henceforth, we put $\zeta \leq \xi \iff \forall n \geq 0 \quad \kappa_n(\zeta) \leq \kappa_n(\xi)$ for $\zeta, \xi \in \overline{\mathcal{M}}$, consistently with the previous definition in case $\zeta, \xi \in \mathcal{M}$. This order is coarser than the prefix ordering on sequences of cliques.

For each $u \in \mathcal{M}$, we define two kinds of *elementary cylinders of base u* :

$$\uparrow u = \{\xi \in \partial\mathcal{M} : u \leq \xi\} \subseteq \partial\mathcal{M}, \quad \uparrow\uparrow u = \{v \in \overline{\mathcal{M}} : u \leq v\} \subseteq \overline{\mathcal{M}}. \quad (1)$$

The set \mathcal{M} being countable, it is equipped with the discrete topology. The set $\overline{\mathcal{M}}$ is a compactification of \mathcal{M} , when equipped with the topology generated by the opens of \mathcal{M} and all cylinders $\uparrow\uparrow u$, for u ranging over \mathcal{M} . This makes $\overline{\mathcal{M}}$ a metrisable compact space [1]. The set $\partial\mathcal{M}$ is a closed subset of $\overline{\mathcal{M}}$. The induced topology on $\partial\mathcal{M}$ is generated by the family of cylinders $\uparrow u$, for u ranging over \mathcal{M} . Finally, both spaces are equipped with their respective Borel σ -algebras, \mathfrak{F} on \mathcal{M} and $\overline{\mathfrak{F}}$ on $\overline{\mathcal{M}}$; the σ -algebra on each space is generated by the corresponding family of cylinders.

Möbius polynomial. Principal root. Sub-uniform measures. We recall [6, 18] the definitions of the *Möbius polynomial* $\mu_{\mathcal{M}}(X)$ and of the *growth series* $G(X)$ associated to \mathcal{M} :

$$\mu_{\mathcal{M}}(X) = \sum_{c \in \mathcal{C}} (-1)^{|c|} X^{|c|}, \quad G(X) = \sum_{u \in \mathcal{M}} X^{|u|} = \sum_{n \geq 0} \lambda_{\mathcal{M}}(n) X^n, \quad (2)$$

where $\lambda_{\mathcal{M}}(n) = \#\{x \in \mathcal{M} : |x| = n\}$. It is known that $G(X)$ is rational, inverse of the Möbius polynomial:

$$G(X) = 1/\mu_{\mathcal{M}}(X).$$

It is also known [14, 7] that $\mu_{\mathcal{M}}(X)$ has a unique root of smallest modulus, say p_0 , which lies in the real interval $(0, 1)$ if $|A| > 1$ (the case $|A| = 1$ is trivial). The root p_0 will be called the *principal root* of $\mu_{\mathcal{M}}$, or simply of \mathcal{M} .

The following result, to be compared with Lemma 1, adapts the so-called Patterson-Sullivan construction from geometric group theory. The compactness of $\overline{\mathcal{M}}$ is an essential ingredient of the proof for the case $p = p_0$, based on classical results from Functional Analysis.

• **Theorem 3**—*For each $p \in (0, p_0]$, where p_0 is the principal root of \mathcal{M} , there exists a unique probability measure ν_p on $(\overline{\mathcal{M}}, \overline{\mathfrak{F}})$ such that $\nu_p(\uparrow x) = p^{|x|}$ holds for all $x \in \mathcal{M}$. On the one hand, if $p < p_0$, then ν_p is concentrated on \mathcal{M} , and is given by:*

$$\forall x \in \mathcal{M} \quad \nu_p(\{x\}) = p^{|x|}/G(p). \quad (3)$$

On the other hand, ν_{p_0} is concentrated on the boundary, hence $\nu_{p_0}(\partial\mathcal{M}) = 1$. In this case, $\nu_{p_0}(\uparrow x) = p_0^{|x|}$ holds for all $x \in \mathcal{M}$.

• **Definition 1**—*The measures ν_p on $\overline{\mathcal{M}}$ described in Th. 3 are called sub-uniform measures of parameter p . The measure ν_{p_0} is called the uniform measure on $\partial\mathcal{M}$.*

The following result relates the uniform measure on the boundary with the sequence $\nu_{\mathcal{M}_k}$ of uniform distributions over the sets $\mathcal{M}_k = \{x \in \mathcal{M} : |x| = k\}$.

• **Theorem 4**—*Let \mathcal{M} be a trace monoid, of principal root p_0 . The sequence of uniform distributions $(\nu_{\mathcal{M}_k})_{k \geq 0}$ converges weakly toward the uniform measure ν_{p_0} on $\partial\mathcal{M}$.*

Anticipating on Th. 5 below, Theorem 4 above has the following concrete consequence. Fix an integer $j \geq 1$, and draw traces of length k uniformly at random, with k arbitrarily large. Then the j first cliques of the trace obtained approximately behave as if they were a Markov chain (C_1, \dots, C_j) ; and the larger k , the better the approximation. Conversely, how this can be exploited for random generation purposes, is the topic of Sect. 4.

Irreducibility and irreducible components. Generators of a trace monoid only have partial commutativity properties. The following definition isolates the parts of the alphabet that enjoy full commutativity.

• **Definition 2**—*Let (A, I) be an independence pair. The associated dependence pair is (A, D) where $D = (A \times A) \setminus I$. The connected components of the graph (A, D) are called the irreducible components of $\mathcal{M} = \mathcal{M}(A, I)$. To each of these irreducible component A' is associated the independence relation $I' = I \cap (A' \times A')$. The corresponding trace monoids $\mathcal{M}' = \mathcal{M}(A', I')$ are called the irreducible components of the trace monoid \mathcal{M} . If (A, D) is connected, then \mathcal{M} is said to be irreducible.*

Direct products of trace monoids are trace monoids themselves. More precisely, the following result holds.

• **Proposition 1**—*Let $\mathcal{M} = \mathcal{M}(A, I)$ be a trace monoid. Then \mathcal{M} is the direct product of its irreducible components. As a measurable space and as a topological space, $\overline{\mathcal{M}}$ is the product of the $\overline{\mathcal{M}'}$, where \mathcal{M}' ranges over the irreducible components of \mathcal{M} . The Möbius polynomial $\mu_{\mathcal{M}}(X)$ is the product of the Möbius polynomials $\mu_{\mathcal{M}'}(X)$, for \mathcal{M}' ranging over the irreducible components of \mathcal{M} .*

The sets $\mathcal{M}_k = \{x \in \mathcal{M} : |x| = k\}$ do not enjoy a product decomposition with respect to irreducible components of \mathcal{M} , hence neither do the uniform distributions $\nu_{\mathcal{M}_k}$ over \mathcal{M}_k . By contrast, sub-uniform measures have a product decomposition, as stated below.

• **Proposition 2**—*Let \mathcal{M} be a trace monoid, of principal root p_0 , and let ν_p be a sub-uniform measure on $\overline{\mathcal{M}}$ with $p \leq p_0$. Then ν_p is the product of measures ν' on each of the $\overline{\mathcal{M}'}$, for \mathcal{M}' ranging over the irreducible components of \mathcal{M} . The measures ν' are all sub-uniform measures on $\overline{\mathcal{M}'}$ of the same parameter p .*

It follows from Prop. 1 that the principal root of a trace monoid \mathcal{M} is the smallest among the principal roots of its irreducible components. As a consequence of Prop. 2, the uniform measure is a product of *sub-uniform* measures ν' over the irreducible components \mathcal{M}' of \mathcal{M} . By Th. 3, each ν' is either concentrated on \mathcal{M}' if the principal root p' of \mathcal{M}' satisfies $p' > p_0$, or concentrated on $\partial\mathcal{M}'$ if $p' = p_0$. Note that at least one of these sub-uniform measures is actually *uniform* on the irreducible component.

Realization of uniform and sub-uniform measures. The characterization of the uniform measure by $\nu(\uparrow x) = p_0^{|x|}$ (see Th. 3) does not provide an obvious recursive procedure for an algorithmic approximation of ν -generated samples on $\partial\mathcal{M}$. Since the uniform measure ν is, according to Prop. 2, a product of sub-uniform measures, it is enough to focus on the algorithmic sampling of *sub-uniform measures on irreducible trace monoids*.

Hence, let \mathcal{M} be an irreducible trace monoid, of principal root p_0 , and let $\overline{\mathcal{M}}$ be equipped with a sub-uniform measure ν_p with $p \leq p_0$. Recall from Th. 3 that ν_p is either concentrated on \mathcal{M} or on $\partial\mathcal{M}$ according to whether $p < p_0$ or $p = p_0$.

Elements of \mathcal{M} are given as finite paths in the graph $(\mathcal{C}, \rightarrow)$, whereas elements of $\partial\mathcal{M}$ are given as infinite paths in $(\mathcal{C}, \rightarrow)$. In order to have a unified presentation of both spaces, we use the following technical trick: instead of considering the graph of *non empty* cliques $(\mathcal{C}, \rightarrow)$, we use the graph of *all* cliques $(\mathcal{C}, \rightarrow)$, including the empty clique. We keep the same definition of the Cartier-Foata relation ' \rightarrow ' (see above). Note that $c \rightarrow 1_{\mathcal{M}}$ then holds for every clique $c \in \mathcal{C}$, whereas $1_{\mathcal{M}} \rightarrow c$ holds if and only if $c = 1_{\mathcal{M}}$. Hence $1_{\mathcal{M}}$ is an absorbing state in $(\mathcal{C}, \rightarrow)$. Any path in $(\mathcal{C}, \rightarrow)$, either finite or infinite, now corresponds to a unique infinite path in $(\mathcal{C}, \rightarrow)$. If the original path $(c_k)_{1 \leq k \leq N}$ is finite, the corresponding infinite path $(c'_k)_{k \geq 1}$ in $(\mathcal{C}, \rightarrow)$ is defined by $c'_k = c_k$ for $1 \leq k \leq N$ and $c'_k = 1_{\mathcal{M}}$ for all $k > N$.

For each trace $\xi \in \overline{\mathcal{M}}$, either finite or infinite, let $(C_k)_{k \geq 1}$ be the infinite sequence of cliques corresponding to the infinite path in $(\mathcal{C}, \rightarrow)$ associated with ξ . The sequence $(C_k)_{k \geq 1}$ is a random sequence of cliques; its characterization under a sub-uniform measure ν_p is the topic of next result.

• **Theorem 5**—*Let \mathcal{M} be an irreducible trace monoid of principal root p_0 . Then, with respect to the sub-uniform measure ν_p on $\overline{\mathcal{M}}$, with $0 < p \leq p_0$, the sequence of random cliques $(C_k)_{k \geq 1}$ is a Markov chain with state space \mathcal{C} .*

Let $g, h : \mathcal{C} \rightarrow \mathcal{R}$ be the functions defined by:

$$h(c) = \sum_{c' \in \mathcal{C} : c' \geq c} (-1)^{|c'| - |c|} p^{|c'|}, \quad g(c) = h(c)/p^{|c|}. \quad (4)$$

Then $(h(c))_{c \in \mathcal{C}}$ is a probability vector over \mathcal{C} , which is the distribution of the initial clique C_1 . This vector is positive on \mathfrak{C} , and $h(1_{\mathcal{M}}) > 0$ if and only if $p < p_0$. The transition matrix of the chain, say $P = (P_{c,c'})_{(c,c') \in \mathcal{C} \times \mathcal{C}}$, is:

$$P_{c,c'} = \begin{cases} 0, & \text{if } c \rightarrow c' \text{ does not hold,} \\ h(c')/g(c), & \text{if } c \rightarrow c' \text{ holds,} \end{cases} \quad (5)$$

with the line $(P_{1_{\mathcal{M}},c'})_{c' \in \mathcal{C}}$ corresponding to the empty clique undefined if $p = p_0$.

Conversely, if $p \leq p_0$, and if $(C_k)_{k \geq 1}$ is a Markov chain on \mathcal{C} if $p < p_0$, respectively on \mathfrak{C} if $p = p_0$, with initial distribution h defined in (4) and with transition matrix P defined in (5), and if $Y_k = C_1 \cdot \dots \cdot C_k$, then $(Y_k)_{k \geq 1}$ converges weakly towards the sub-uniform measure ν_p . Furthermore, the law of the random trace $C_1 \cdot C_2 \cdot \dots = \bigvee_{k \geq 1} Y_k \in \overline{\mathcal{M}}$ is the probability measure ν_p on $\overline{\mathcal{M}}$.

Theorem 5 for $p = p_0$ already appears in [2]. Note: the function $h : \mathcal{C} \rightarrow \mathcal{R}$ defined in (4) is the *Möbius transform* in the sense of Rota [17, 3] of the function $f : c \in \mathcal{C} \mapsto p^{|c|}$; see [2] for more emphasis on this point of view.

As expected, we recover the results of § 2 in the case of two commuting alphabets A and B with $|A| > |B|$. Indeed, by Prop. 2 and Th. 5, the cliques $(C_k)_{k \geq 1}$ form a product of two Markov chains: one on A (*non empty cliques of A^**) and the other one on $B \cup \{1_{B^*}\}$ (*cliques of B^* , including the empty one*).

4—Uniform Generation of Finite Traces

We have introduced in Def. 1 a notion of uniform measure on the boundary of a trace monoid. This measure is characterized by its values on cylinders in Th. 3, as the weak limit of uniform distributions in Prop. 2, and through the associated Cartier-Foata probabilistic process in Th. 5.

Because of the existence of the Cartier-Foata normal form of traces, the combinatorics of a trace monoid is entirely contained in the Cartier-Foata automaton, either $(\mathcal{C}, \rightarrow)$ or $(\mathfrak{C}, \rightarrow)$. Looking at the Cartier-Foata automaton, say $(\mathfrak{C}, \rightarrow)$ on non empty-cliques, as generating a sub-shift of finite type, it is interesting to investigate the associated notion of uniform measure ‘à la Parry’ [16, 13, 15], and to compare it with the uniform measure on the boundary previously introduced. This comparison between the two notions of uniform measures will enlighten the forthcoming discussion on uniform generation of finite traces.

Uniform measure on the boundary versus Parry measure. The Parry measure associated with an *irreducible* sub-shift of finite type is formally defined as the unique measure of maximal entropy on *bi-infinite* admissible sequences of states of the sub-shift. It corresponds intuitively to the “uniform measure” on such bi-infinite paths (see, e.g., [15]).

The Parry measure is only defined for *irreducible* sub-shifts for good reasons. Indeed, if a sub-shift has, say, two parts X and Y , with Y an irreducible component and such that going from X to Y is possible but not the other way around as in Fig. 2-(a), then one cannot define a “uniform measure” on bi-infinite paths (it should put mass on paths spending an infinite amount of

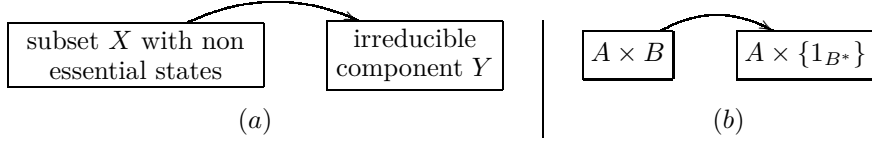


Figure 2: (a) *Illustration of a reducible system* (b) *Cartier-Foata automaton on non-empty cliques of $A^* \times B^*$ generating the uniform measure on $\partial(A^* \times B^*)$ if $|A| > |B|$*

time both in X and Y and be stationary, which is impossible). On the other hand, considering a uniform measure on one-sided infinite sequences on such a compound system makes perfect sense. This is the case, for instance, of the Cartier-Foata sub-shift associated to the reducible trace monoids $A^* \times B^*$ with $|A| > |B|$ studied in § 2: see Fig. 2–(b).

For a general trace monoid \mathcal{M} , the associated sub-shift $(\mathfrak{C}, \rightarrow)$ is irreducible if and only if the monoid \mathcal{M} is irreducible in the sense of Def. 2 (a well-known result: see for instance [14, Lemma 3.2]). Therefore the comparison between the uniform measure on the boundary, and the Parry measure, only makes sense in this case.

Hence, let \mathcal{M} be an irreducible trace monoid, of principal root p_0 . In order to take into account the length of cliques in the construction of the Parry measure, we consider the weighted incidence matrix $B = (B_{x,y})_{(x,y) \in \mathfrak{C} \times \mathfrak{C}}$ defined by $B_{x,y} = p_0^{|y|}$ if $x \rightarrow y$ holds and by $B_{x,y} = 0$ if $x \rightarrow y$ does not hold.

• **Lemma 2**—*The non-negative matrix B has spectral radius 1. The vector $g = (g(c))_{c \in \mathfrak{C}}$ defined by $g(c) = \sum_{c' \in \mathfrak{C} : c \rightarrow c'} h(c')$ for $c \in \mathfrak{C}$, where h has been defined in (4), is B -invariant on the right: $Bg = g$.*

Define the matrix $C = (C_{c,c'})_{(c,c') \in \mathfrak{C} \times \mathfrak{C}}$ by:

$$\forall c, c' \in \mathfrak{C} \quad C_{c,c'} = B_{c,c'} g(c') / g(c). \quad (6)$$

Since g is right invariant for B , it follows that C is stochastic. Classically, the Parry measure on bi-infinite paths in $(\mathfrak{C}, \rightarrow)$ is the stationary Markovian measure of transition matrix C .

• **Proposition 3**—*The matrix C defined in (6) coincides with the transition matrix P defined in Theorem 5 for $p = p_0$, and restricted to $\mathfrak{C} \times \mathfrak{C}$.*

Proposition 3 asserts that the Markov chain associated with the Parry measure has the same transition matrix as the probabilistic process on non-empty cliques generated by the uniform measure on the boundary. But the Parry measure is stationary whereas the uniform measure ν is not. Indeed, the initial distribution of the Markov measure ν is $h : \mathfrak{C} \rightarrow \mathcal{R}$, which does *not* coincide with the stationary measure of the chain (except in the trivial case of a free monoid).

To summarize: the notion of uniform measure on the boundary is adapted to one-sided infinite heaps, independently of the irreducibility of the trace monoid under consideration. If the monoid is irreducible, there is a notion of uniform measure on two-sided infinite heaps, which correspond to a weighted Parry

measure. Considering the projection of this Parry measure to one-sided infinite heaps, and conditionally on a given initial clique, it coincides with the uniform measure at infinity since they share the same transition matrix. But the two measures globally differ since their initial measures differ.

Uniform generation of finite traces, 0. The Parry measure is a standard tool for a special type of uniform generation. Indeed, it provides an algorithmic way of sampling finite sequences of a fixed length k , and uniformly *if the first and the last letters of the sequence are given*. In our framework, besides the fact that the Parry measure is only defined for an irreducible trace monoid, it also misses the primary target of generating finite traces of a given length k among *all traces* of length k .

Uniform generation of finite traces, 1. Consider the problem, given a fixed integer $k > 1$ and a trace monoid $\mathcal{M} = \mathcal{M}(A, I)$, of designing a randomized algorithm which produces a trace $x \in \mathcal{M}$ of length k , uniformly among traces of length k . Sub-uniform measures on the trace monoid \mathcal{M} allow to adapt to our framework the technique of Boltzmann samplers [10] for solving this problem.

Consider a parameter $p \in (0, p_0)$, where p_0 is the principal root of \mathcal{M} , and let $\xi \in \mathcal{M}$ be sampled according to the sub-uniform measure ν_p . We have indeed $|\xi| < \infty$ with probability 1 by Th. 3. Furthermore, Prop. 2 shows that ν_p decomposes as a product of sub-uniform measures of the same parameter p , over the irreducible components of \mathcal{M} . For each component, sampling is done through usual Markov chain generation techniques since both the initial measure and the transition matrix of the chain of cliques are explicitly known by Th. 5.

The algorithm is then the following: if $|\xi| = k$, then keep ξ ; otherwise, reject ξ and sample another trace. This eventually produces a random trace of length k , uniformly distributed in \mathcal{M}_k ; since ν_p is a weighted sum of all $\nu_{\mathcal{M}_k}$, as shown by the expression (3).

As usual, the optimal parameter p , for which the rejection probability is the lowest, is such that: $\mathbb{E}_{\nu_p}|\xi| = k$, where $\mathbb{E}_{\nu_p}(\cdot)$ denotes the expectation with respect to ν_p . Ordinary computations show that $\mathbb{E}_{\nu_p}|\xi|$ is related to the derivative of the growth function by $\mathbb{E}_{\nu_p}|\xi| = pG'(p)/G(p) = -p\mu'_{\mathcal{M}}(p)/\mu_{\mathcal{M}}(p)$; providing an explicit equation

$$k\mu_{\mathcal{M}}(p) + p\mu'_{\mathcal{M}}(p) = 0 ,$$

to be numerically solved in p .

Unfortunately, the rejection probability approaches 1 exponentially fast as k increases, making the algorithm less and less efficient. A standard way to overcome this difficulty would be to consider *approximate sampling* [10], consisting in sampling traces of length approximately k .

Uniform generation of finite traces, 2: evaluating an average cost. Uniform generation is often done in order to evaluate the expected value of a cost function. For this purpose, a more direct approach in our framework is based on an exact integration formula given in Th. 6 below.

Let $\phi : \mathcal{M}_k \rightarrow \mathcal{R}$ be a cost function, and consider the problem of evaluating the expectation $\mathbb{E}_{\nu_{\mathcal{M}_k}}(\phi)$, for a fixed integer k . For each integer $k \geq 0$, let:

$$\mathcal{M}_k = \{x \in \mathcal{M} : |x| = k\}, \quad \lambda_{\mathcal{M}}(k) = \#\mathcal{M}_k, \quad \mathcal{M}_{(k)} = \{x \in \mathcal{M} : \tau(x) = k\}.$$

To each function $\phi : \mathcal{M}_k \rightarrow \mathcal{R}$ defined on traces of *length* k , we associate a function $\bar{\phi} : \mathcal{M}_{(k)} \rightarrow \mathcal{R}$ defined on traces of *height* k , as follows:

$$\forall x \in \mathcal{M}_{(k)} \quad \bar{\phi}(x) = \sum_{y \in \mathcal{M}_k : y \leq x} \phi(y). \quad (7)$$

• **Theorem 6**—Let $\bar{\phi} : \mathcal{M}_{(k)} \rightarrow \mathcal{R}$ be defined as in (7). Then the following equality holds between the expectation with respect to the uniform distribution $\nu_{\mathcal{M}_k}$ on \mathcal{M}_k on the one hand, and the expectation with respect to the uniform measure ν on $\partial\mathcal{M}$ on the other hand (whether \mathcal{M} is irreducible or not):

$$\mathbb{E}_{\nu_{\mathcal{M}_k}} \phi = (p_0^k \cdot \lambda_{\mathcal{M}}(k))^{-1} \cdot \mathbb{E}_{\nu} \bar{\phi}(C_1 \cdot \dots \cdot C_k). \quad (8)$$

The generation of $(C_k)_{k \geq 1}$ enables us to evaluate $\mathbb{E}_{\nu} \bar{\phi}(C_1 \cdot \dots \cdot C_k)$ for any integer k , provided the function $\bar{\phi}$ can be efficiently computed. In turn, this directly depends on the numbers $\theta_k(x) = \#\{y \in \mathcal{M}_k : y \leq x\}$ of terms in the sum (7) defining $\bar{\phi}(x)$. The numbers $\theta_k(x)$ might be arbitrary large; for instance $\theta_k((a \cdot b)^k) = k + 1$ for $(a, b) \in I$. However we have the following result.

• **Lemma 3**—Assume that \mathcal{M} is irreducible. Then, there exists $C > 0$ such that:

$$\mathbb{E}_{\nu} \theta_k(C_1 \cdot \dots \cdot C_k) \leq C.$$

To see this, apply (8) to the constant function $\phi = 1$ on \mathcal{M}_k , whose associated function is $\bar{\phi} = \theta_k$ on $\mathcal{M}_{(k)}$, to obtain:

$$\mathbb{E}_{\nu} \theta_k(C_1 \cdot \dots \cdot C_k) = p_0^k \cdot \lambda_{\mathcal{M}}(k). \quad (9)$$

The terms $\lambda_{\mathcal{M}}(k)$, coefficients of the growth series $G(X) = 1/\mu_{\mathcal{M}}(X)$, are asymptotically equivalent to Cp_0^{-k} for some constants $C > 0$ if \mathcal{M} is irreducible [14]. The result in Lemma 3 follows.

Applying usual techniques [4] to specifically retrieve all traces $y \leq x$ of length $k = \tau(x)$ is feasible in time $O(k)$ in average and allows to compute $\bar{\phi}(x)$, and consequently to estimate the expectation $\mathbb{E}_{\nu} \bar{\phi}(C_1 \cdot \dots \cdot C_k)$ via Markov chain sampling and a Monte-Carlo algorithm.

By (9), applying the same estimation technique to the function $\phi = 1$ yields an estimate for the normalization factor $p_0^k \cdot \lambda_{\mathcal{M}}(k)$. In passing, this also yields a Monte-Carlo estimate for the number $\lambda_{\mathcal{M}}(k)$. All together, we are thus able to estimate with an arbitrary precision both terms in the right hand member of (8), hence yielding an accurate estimation of $\mathbb{E}_{\nu_{\mathcal{M}_k}} \phi$.

To summarize: generating the first k layers of traces under the uniform measure on the boundary allows to compute the *expectation* of an arbitrary computable cost function $\phi : \mathcal{M}_k \rightarrow \mathbb{R}$, if \mathcal{M} is irreducible. The same applies at the cost of a greater complexity if \mathcal{M} is not irreducible.

References

- [1] S. Abbes and K. Keimel. “Projective topology on bifinite domains and applications”. In: *Theoret. Comput. Sci.* 365.3 (2006), pp. 171–183.
- [2] S. Abbes and J. Mairesse. “Uniform and Bernoulli measures on the boundary of trace monoids”. In: *J. Combin. Theory Ser. A* 135 (2015), pp. 201–236.
- [3] M. Aigner. *A Course in Enumeration*. Springer, 2007.
- [4] A. Bertoni et al. “Counting techniques for inclusion, equivalence and membership problems”. In: *The Book of Traces*. World Scientific, 1994, pp. 131–163.
- [5] O. Bodini, A. Genitrini, and F. Peschanski. “Enumeration and random generation of concurrent computations”. In: *Proc. of AofA’12*. DMTCS, 2012, pp. 83–96.
- [6] P. Cartier and D. Foata. *Problèmes combinatoires de commutation et réarrangements*. Vol. 85. Lecture Notes in Math. Springer, 1969.
- [7] P. Csikvári. “Note on the smallest root of the independence polynomial”. In: *Combin. Probab. Comput.* 22.1 (2013), pp. 1–8.
- [8] V. Diekert. *Combinatorics on Traces*. Vol. 454. LNCS. Springer, 1990.
- [9] V. Diekert and G. Rozenberg, eds. *The Book of Traces*. World Scientific, 1995.
- [10] P. Duchon et al. “Boltzmann samplers for the random generation of combinatorial structures”. In: *Combin. Probab. Comput.* 13 (2004), pp. 577–625.
- [11] P. Flajolet, P. Zimmermann, and B. Van Cutsem. “Calculus for the random generation of labelled combinatorial structures”. In: *Theoret. Comp. Sci.* 218.2 (1994), pp. 233–248.
- [12] M. Jerrum. *Counting, sampling and integrating: algorithms and complexity*. Springer, 2013.
- [13] B.P. Kitchens. *Symbolic dynamics. One-sided, two-sided and countable state Markov shifts*. Springer, 1998.
- [14] D. Krob, J. Mairesse, and I. Michos. “Computing the average parallelism in trace monoids”. In: *Discrete Math.* 273 (2003), pp. 131–162.
- [15] D. Lind and B. Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, 1995.
- [16] W. Parry. “Intrinsic Markov chains”. In: *Trans. Amer. Math. Soc.* 112.1 (1964), pp. 55–66.
- [17] G.-C. Rota. “On the foundations of combinatorial theory I. Theory of Möbius functions”. In: *Z. Wahrscheinlichkeitstheorie* 2 (1964), pp. 340–368.
- [18] X. Viennot. “Heaps of pieces, I : basic definitions and combinatorial lemmas”. In: *Combinatoire énumérative*. Vol. 1234. Lecture Notes in Math. Springer, 1986, pp. 321–350.

Appendix: proofs

Proof of Theorem 3. We first recall that two traces $x, y \in \mathcal{M}$ having a common upper bound have a least upper bound $x \vee y$, in which case x and y are said to be *compatible*. Therefore:

$$\uparrow x \cap \uparrow y = \begin{cases} \emptyset, & \text{if } x, y \text{ are not compatible,} \\ \uparrow (x \vee y), & \text{if } x, y \text{ are compatible.} \end{cases}$$

It follows that the collection of elementary cylinders $\uparrow x$, to which is added the empty set, is stable by finite intersections. They form thus a π -system. It follows that a measure ν on $(\overline{\mathcal{M}}, \overline{\mathfrak{F}})$ is entirely determined by its values on all elementary cylinders $\uparrow x$, for $x \in \mathcal{M}$. The uniqueness stated in the theorem follows.

We also recall that, since the growth series $G(X)$ has positive coefficients on the one hand, and since the formal equality $G(X) = 1/\mu_{\mathcal{M}}(X)$ holds on the other hand, the radius of convergence of the power series $G(z)$ is exactly p_0 .

Let ν_p be the probability measure on $\overline{\mathcal{M}}$ defined, for $p \in [0, p_0)$, by:

$$\nu_p = \frac{1}{G(p)} \sum_{y \in \mathcal{M}} p^{|y|} \delta_y,$$

where δ_y denotes the Dirac measure concentrated on $\{y\}$. The measure ν_p is well defined for $0 \leq p < p_0$ since $G(p) < +\infty$, as recalled above.

We prove that $\nu_p(\uparrow x) = p^{|x|}$ for all $x \in \mathcal{M}$. Observe that the mapping $y \in \mathcal{M} \mapsto x \cdot y$ is a bijection onto $\uparrow x \cap \mathcal{M}$. Since ν_p is concentrated on \mathcal{M} , we compute for any $x \in \mathcal{M}$:

$$\begin{aligned} \nu_p(\uparrow x) &= \nu_p(\uparrow x \cap \mathcal{M}) \\ &= \frac{1}{G(p)} \sum_{y \in \mathcal{M} : y \geq x} p^{|y|} \\ &= \frac{1}{G(p)} \sum_{y \in \mathcal{M}} p^{|x \cdot y|} \\ &= p^{|x|} \frac{1}{G(p)} \sum_{y \in \mathcal{M}} p^{|y|} \\ &= p^{|x|} \end{aligned}$$

This shows that $\nu_p(\uparrow x) = p^{|x|}$ holds for all $x \in \mathcal{M}$.

For $p = p_0$, we adapt the construction of the so-called Patterson-Sullivan measure. Consider any weak limit, say ν_{p_0} , of $(\nu_p)_p$ by letting $p \rightarrow p_0$. Such a limit exists since $\overline{\mathcal{M}}$ is compact and therefore any sequence of probabilities on $\overline{\mathcal{M}}$ has a weakly convergent subsequence. In $\overline{\mathcal{M}}$, any cylinder $\uparrow x$ is both open and closed, therefore its topological boundary is empty, and therefore has

null ν_{p_0} -measure. By the Porte-manteau theorem (see for instance Billingsley's *Convergence of probability measures*), we have thus:

$$\nu_{p_0}(\uparrow x) = \lim_{p \rightarrow p_0} \nu_p(\uparrow x) = \lim_{p \rightarrow p_0} p^{|x|} = p_0^{|x|}. \quad (10)$$

For the same reasons, we have for every $x \in \mathcal{M}$:

$$\begin{aligned} \nu_{p_0}(\{x\}) &= \lim_{p \rightarrow p_0} \nu_p(\{x\}) \\ &= \lim_{p \rightarrow p_0} \frac{p^{|x|}}{G(p)} \\ &= 0, \end{aligned}$$

since $G(p_0) = +\infty$. Since \mathcal{M} is countable, it follows that $\nu_{p_0}(\mathcal{M}) = 0$, and thus ν_{p_0} is concentrated on the boundary. Finally, since $\uparrow x = \uparrow x \cap \partial\mathcal{M}$, and using that $\nu_{p_0}(\mathcal{M}) = 0$, we obtain:

$$\nu_{p_0}(\uparrow x) = \nu_{p_0}(\uparrow x \cap \partial\mathcal{M}) = \nu_{p_0}(\uparrow x) = p_0^{|x|}.$$

This completes the proof of Theorem 3.

Proof of Theorem 4. Since $\overline{\mathcal{M}}$ is compact, to show that $(\nu_{\mathcal{M}_k})_k$ converges toward ν_{p_0} , it suffices to show that ν_{p_0} is the weak limit of any weakly convergent subsequence of $(\nu_{\mathcal{M}_k})_k$. Let ν be the weak limit of a weakly convergent subsequence $(\nu_{\mathcal{M}_{k_j}})_j$. Using the estimate $\#\mathcal{M}_k \sim_{k \rightarrow \infty} Ck^N p_0^{-k}$ (see [14]), we have for all $x \in \mathcal{M}$ and for all j large enough:

$$\begin{aligned} \nu_{\mathcal{M}_{k_j}}(\uparrow x) &= \frac{1}{\#\mathcal{M}_{k_j}} \#\{y \in \mathcal{M} : |y| = k_j \wedge y \geq x\} \\ &= \frac{1}{\#\mathcal{M}_{k_j}} \#\{y \in \mathcal{M} : |y| = k_j - |x|\} \\ &\sim_{j \rightarrow \infty} \frac{1}{C(k_j)^N p_0^{-k_j}} C(k_j - |x|)^N p_0^{|x| - k_j} \\ &\rightarrow_{j \rightarrow \infty} p_0^{|x|}. \end{aligned}$$

Using the Porte-manteau theorem as in the proof of Th. 3, we have thus:

$$\nu(\uparrow x) = \lim_{j \rightarrow \infty} \nu_{\mathcal{M}_{k_j}}(\uparrow x) = p_0^{|x|},$$

and therefore $\nu = \nu_{p_0}$, using as above that measures are entirely determined by their values on elementary cylinders. This shows that $(\nu_{\mathcal{M}_k})_k$ converges toward ν_{p_0} . The proof of Theorem 4 is complete.

Proof of Proposition 1. For simplicity, assume that \mathcal{M} has exactly two irreducible components \mathcal{M}_1 and \mathcal{M}_2 , and let A_1 and A_2 be the corresponding irreducible components of A .

Let $i_1 : \mathcal{M}_1 \rightarrow \mathcal{M}$ and $i_2 : \mathcal{M}_2 \rightarrow \mathcal{M}$ be the natural injections, and let $f : \mathcal{M}_1 \times \mathcal{M}_2 \rightarrow \mathcal{M}$ be defined by $f(x_1, x_2) = i_1(x_1) \cdot i_2(x_2) = i_2(x_2) \cdot i_1(x_1)$.

Then it is clear that f is an isomorphism. The natural morphism $\pi_1 : \mathcal{M} \rightarrow \mathcal{M}_1$ for instance, is entirely determined by:

$$\forall x \in A_1 \cup A_2 \quad \pi_1(x) = \begin{cases} 0, & \text{if } x \in A_2, \\ x, & \text{if } x \in A_1. \end{cases}$$

We extend f to $\overline{\mathcal{M}}_1 \times \overline{\mathcal{M}}_2 \rightarrow \overline{\mathcal{M}}$ as follows. Let us first identify \mathcal{M}_1 and \mathcal{M}_2 with sub-monoids of \mathcal{M} , through the natural injections i_1 and i_2 . The two following properties are obvious:

- (a) $\forall x_1 \in \mathcal{M}_1 \quad \forall x_2 \in \mathcal{M}_2 \quad x_1 \cdot x_2 = x_2 \cdot x_1$
- (b) $\forall u, v, w \in \mathcal{M} \quad u \geq v \implies w \cdot u \geq w \cdot v$

Now let $(x, y) \in \overline{\mathcal{M}}_1 \times \overline{\mathcal{M}}_2$. Let $(c_n)_{n \geq 1}$ and $(d_n)_{n \geq 1}$ be the Cartier-Foata sequence of cliques of \mathcal{M}_1 and of \mathcal{M}_2 associated to x and y respectively. It is not difficult to see that, by putting $x_n = c_1 \cdot \dots \cdot c_n$ and $y_n = d_1 \cdot \dots \cdot d_n$, one has:

$$\bigvee_{n \geq 1} x_n = x \quad \text{in } \overline{\mathcal{M}}_1, \quad \bigvee_{n \geq 1} y_n = y \quad \text{in } \overline{\mathcal{M}}_2.$$

Then observe that the sequence $(x_n \cdot y_n)_{n \geq 1}$ is non decreasing in \mathcal{M} . Indeed:

$$\begin{aligned} x_{n+1} \cdot y_{n+1} &\geq x_{n+1} \cdot y_n && \text{by (b) and since } y_{n+1} \geq y_n \\ &= y_n \cdot x_{n+1} && \text{by (a)} \\ &\geq y_n \cdot x_n && \text{by (b) and since } x_{n+1} \geq x_n \\ &= x_n \cdot y_n && \text{by (a).} \end{aligned}$$

Since $\overline{\mathcal{M}}$ is complete w.r.t. the least upper bound of non decreasing sequences (see [2, § 2.1]), we define $f(x, y) \in \overline{\mathcal{M}}$ by:

$$f(x, y) = \bigvee_{n \geq 1} (x_n \cdot y_n) \quad \text{in } \overline{\mathcal{M}}.$$

It is then routine to see that f thus defined is a bijection $\overline{\mathcal{M}}_1 \times \overline{\mathcal{M}}_2 \rightarrow \overline{\mathcal{M}}$.

Using the natural morphisms $\pi_1 : \mathcal{M} \rightarrow \mathcal{M}_1$ and $\pi_2 : \mathcal{M} \rightarrow \mathcal{M}_2$, we have:

$$\begin{aligned} \forall x \in \mathcal{M} \quad f^{-1}(\uparrow x) &= \uparrow(\pi_1(x)) \times \uparrow(\pi_2(x)), \\ \forall (x_1, x_2) \in \mathcal{M}_1 \times \mathcal{M}_2 \quad f(\uparrow x_1 \times \uparrow x_2) &= \uparrow(x_1 \cdot x_2). \end{aligned}$$

This shows that f is a homeomorphism, hence *a fortiori* bi-measurable.

Finally, let \mathcal{C} denote the set of cliques of \mathcal{M} , and let $\mathcal{C}_1, \mathcal{C}_2$ denote the sets of cliques of \mathcal{M}_1 and \mathcal{M}_2 respectively. The isomorphism $f : \mathcal{M}_1 \times \mathcal{M}_2 \rightarrow \mathcal{M}$ induces by restriction a bijection $\mathcal{C}_1 \times \mathcal{C}_2 \rightarrow \mathcal{C}$, from which follows the product decomposition $\mu_{\mathcal{M}} = \mu_{\mathcal{M}_1} \times \mu_{\mathcal{M}_2}$. The proof of Proposition 1 is complete.

Proof of Proposition 2. Assume for simplicity that \mathcal{M} has two irreducible components, say \mathcal{M}_1 and \mathcal{M}_2 . For $x \in \mathcal{M}$, let $x_1 = \pi_1(x) \in \mathcal{M}_1$ and $x_2 = \pi_2(x) \in \mathcal{M}_2$ be the components of x in \mathcal{M}_1 and in \mathcal{M}_2 . Then we have:

$$\nu_p(\uparrow x) = p^{|x|} = p^{|x_1|} \times p^{|x_2|} = \nu'_p(\uparrow x_1) \times \nu''_p(\uparrow x_2),$$

where ν'_p is the sub-uniform measure of parameter p on $\overline{\mathcal{M}}_1$, and ν''_p is the sub-uniform measure of parameter p on $\overline{\mathcal{M}}_2$. Using again that the collection of elementary cylinders is a π -system, this is enough to conclude that $\nu_p = \nu'_p \otimes \nu''_p$. The proof of Proposition 2 is complete.

Proof of Theorem 5. Our proof extends the proofs of [2, Th. 2.4, Th. 2.5] by taking into account the possible presence of the empty clique in the Cartier-Foata decomposition of traces. It also makes a specific use of the existence of the uniform measure obtained through the construction of Theorem 3. Let p be a real number such that $p \leq p_0$.

For two cliques $c, c' \in \mathcal{C}$, let us write $c \parallel c'$ whenever $c \times c' \subseteq I$, which is equivalent to saying both $c \cap c' = \emptyset$ and $c \cdot c' \in \mathcal{C}$.

Let us denote by C_1, \dots, C_k the first k cliques in the Cartier-Foata decomposition of a random trace $\xi \in \overline{\mathcal{M}}$. Let $c_1 \rightarrow \dots \rightarrow c_k$ be a Cartier-Foata sequence of cliques, $c_i \in \mathcal{C}$ for all $i \in \{1, \dots, k\}$, and put $x = c_1 \cdot \dots \cdot c_{k-1}$. Then we have:

$$\{\xi \in \overline{\mathcal{M}} : C_1 = c_1, \dots, C_k = c_k\} = \uparrow(x \cdot c_k) \setminus \bigcup_{c \in \mathcal{C} : c > c_k} \uparrow(x \cdot c) \quad (11)$$

Let (a_1, \dots, a_r) be an enumeration of the elements a of the alphabet A such that $a \notin c_k$ and $c_k \cup \{a\} \in \mathcal{C}$, or equivalently with the above notation, those $a \in A$ such that $a \parallel c_k$. Then (11) rewrites as:

$$\{\xi \in \overline{\mathcal{M}} : C_1 = c_1, \dots, C_k = c_k\} = \uparrow(x \cdot c_k) \setminus \bigcup_{j=1}^r \uparrow(x \cdot c_k \cdot a_j)$$

Passing to the probabilities on both sides yields:

$$\nu_p(C_1 = c_1, \dots, C_k = c_k) = \nu_p(\uparrow(x \cdot c_k)) - \underbrace{\nu_p\left(\bigcup_{j=1}^r \uparrow(x \cdot c_k \cdot a_j)\right)}_R, \quad (12)$$

since the union on the right side member is included in $\uparrow(x \cdot c_k)$. We evaluate the term R in (12) using Poincaré inclusion-exclusion principle:

$$\begin{aligned} R &= \sum_{j=1}^r (-1)^{j+1} \sum_{1 \leq l_1 < \dots < l_j \leq r} \nu_p(\uparrow(x \cdot c_k \cdot a_{l_1}) \cap \dots \cap \uparrow(x \cdot c_k \cdot a_{l_j})) \\ &= \sum_{j=1}^r (-1)^{j+1} \sum_{c' \in \mathcal{C} : c_k \parallel c' \wedge |c'|=j} \nu_p(\uparrow(x \cdot c_k \cdot c')) \\ &= \sum_{c' \in \mathcal{C} : c_k \parallel c'} (-1)^{|c'|+1} \nu_p(\uparrow(x \cdot c_k \cdot c')) \\ &= \sum_{\delta \in \mathcal{C} : \delta > c_k} (-1)^{|\delta|-|c_k|+1} \nu_p(\uparrow(x \cdot \delta)) \quad \text{with the change of variable } \delta = c_k \cdot c' \end{aligned}$$

Returning to (12), we get:

$$\begin{aligned} \nu_p(C_1 = c_1, \dots, C_k = c_k) &= p^{|x|+|c_k|} + \sum_{\delta \in \mathcal{C} : \delta > c_k} (-1)^{|\delta|-|c_k|} p^{|x|+|\delta|} \\ &= p^{|x|} \left(\sum_{\delta \in \mathcal{C} : \delta \geq c_k} (-1)^{|\delta|-|c_k|} p^\delta \right). \end{aligned}$$

From this we deduce the following formula:

$$\nu_p(C_1 = c_1, \dots, C_k = c_k) = p^{|x|} h(c_k), \quad \text{with } x = c_1 \cdot \dots \cdot c_{k-1}, \quad (13)$$

if $c_1 \rightarrow \dots \rightarrow c_k$ holds, and where $h : \mathcal{C} \rightarrow \mathcal{R}$ is defined by (4). In particular for $k = 1$, we get:

$$\forall c_1 \in \mathcal{C} \quad \nu_p(C_1 = c_1) = h(c_1),$$

which proves at once that $(h(c))_{c \in \mathcal{C}}$ is a probability vector, and that it is the distribution of the first clique C_1 under ν_p .

Let us prove that h is non zero on \mathfrak{C} . Since \mathcal{M} is irreducible, the graph $(\mathfrak{C}, \rightarrow)$ of non empty cliques is strongly connected (a well known result, see a proof in [14]). Let $c \in \mathfrak{C}$, and let $c' \in \mathfrak{C}$ be maximal in (\mathcal{C}, \leq) . Let c_1, \dots, c_n be $n \geq 2$ non empty cliques such that $c_1 \rightarrow \dots \rightarrow c_n$ holds, and $c_1 = c$ and $c_n = c'$. Then, by (13), we have:

$$\begin{aligned} \nu_p(C_1 = c_1, \dots, C_n = c_n) &= p^{|c_1| + \dots + |c_{n-1}|} h(c_n) \\ &= p^{|c_1| + \dots + |c_{n-1}|} p^{|c_n|}, \end{aligned}$$

since $h(c_n) = p^{|c_n|}$, by the maximality of c_n . In particular:

$$\nu_p(C_1 = c_1) \geq \nu_p(C_1 = c_1, \dots, C_n = c_n) \neq 0.$$

But we also have $\nu_p(C_1 = c_1) = h(c_1)$, and thus $h(c_1) \neq 0$, which was to be shown. The value $h(1_{\mathcal{M}})$ coincides with $h(1_{\mathcal{M}}) = \mu_{\mathcal{M}}(p)$. Since p_0 is the root of smallest modulus of $\mu_{\mathcal{M}}$, and since $p \leq p_0$, it follows that $h(1_{\mathcal{M}}) = 0$ if and only if $p = p_0$.

We now come to the proof that $(C_k)_{k \geq 1}$ is a Markov chain, and to the computation of its transition matrix P . If $p = p_0$, then this is the result of [2, Th. 2.5]. The identity of the transition matrices given in the present statement on the one hand, and in [2, Th. 2.5] on the other hand, follows from [2, Prop. 4.12]. Hence, assume that $p < p_0$.

If $c_1 \rightarrow \dots \rightarrow c_k$ holds, with all $c_j \neq 0$, then the expression (13) combined with the fact that $h \neq 0$ on \mathfrak{C} , implies that $\nu_p(C_1 = c_1, \dots, C_{k-1} = c_{k-1}) \neq 0$. Henceforth the following conditional probability is well defined:

$$\nu_p(C_k = c_k | C_1 = c_1, \dots, C_{k-1} = c_{k-1}) = \frac{p^{|c_1| + \dots + |c_{k-1}|} h(c_k)}{p^{|c_1| + \dots + |c_{k-2}|} h(c_{k-1})} \quad (14)$$

$$= \frac{p^{|c_{k-1}|} h(c_k)}{h(c_{k-1})} \quad (15)$$

In case one of the c_j is the empty clique, then the cliques c_{j+1}, \dots, c_k must also be empty since we assume that $c_1 \rightarrow \dots \rightarrow c_k$ holds, and thus:

$$\nu_p(C_k = 1_{\mathcal{M}} | C_1 = c_1, \dots, C_{k-1} = c_{k-1}) = 1 \quad \text{if one } c_j \text{ with } j < k \text{ is } 1_{\mathcal{M}}. \quad (16)$$

Since $c_{k-1} = c_k = 1_{\mathcal{M}}$, the right member of (15) evaluates to 1 as well in this case. Hence (15) is valid in all cases if $c_1 \rightarrow \dots \rightarrow c_k$ holds. Since the right

member of (15) only depends on c_{k-1}, c_k on the one hand, and since on the other hand it is clear that we have:

$$\nu_p(C_1 = c_1, \dots, C_k = c_k) = 0, \quad \text{if } c_1 \rightarrow \dots \rightarrow c_k \text{ does not hold,}$$

we conclude that $(C_k)_{k \geq 1}$ is indeed a Markov chain with the transition matrix described in the statement of the theorem.

For the proof of the converse part of the statement, we proceed in four steps. Consider the two following claims:

1. The vector $(h(c))_{c \in \mathcal{C}}$ is a probability vector.
2. The matrix P is stochastic.

Since we already know the existence of the measure ν_p by Theorem 3, the results already shown so far in the proof make both points 1–2 immediate. They follow from the mere existence of the Markov chain $(C_k)_{k \geq 1}$ previously defined under the measure ν_p on $\overline{\mathcal{M}}$, since h is the distribution of C_1 and P is the transition matrix of the chain.

For the next claim, we introduce new notations in order to avoid confusion with the Markov chain $(C_k)_{k \geq 1}$ previously defined.

3. Let $(C'_k)_{k \geq 1}$ be a Markov chain on \mathcal{C} with initial distribution h and transition matrix P , and let $Y'_k = C'_1 \dots C'_k$. Then the law of $\bigvee_{k \geq 1} Y'_k \in \overline{\mathcal{M}}$ is ν_p .

Let $(\Omega, \mathfrak{G}, \mathbb{P})$ be the sample space on which the Markov chain $(C'_k)_{k \geq 1}$ is defined, and put $\xi' = \bigvee_{k \geq 1} Y'_k$ (see the proof of Prop. 1 above for the existence of the least upper bound in $\overline{\mathcal{M}}$). Let also $\xi \in \overline{\mathcal{M}}$ be the canonical random variable defined on $\overline{\mathcal{M}}$ with law ν_p . Then we have, for every sequence (c_1, \dots, c_k) of cliques:

$$\mathbb{P}(C'_1 = c_1, \dots, C'_k = c_k) = \nu_p(C_1 = c_1, \dots, C_k = c_k),$$

since $(C_k)_{k \geq 1}$ and $(C'_k)_{k \geq 1}$ have same initial distribution and same transition matrix. Therefore, for every $x \in \mathcal{M}$ and for $k = \tau(x)$, we have:

$$\mathbb{P}(\xi' \geq x) = \mathbb{P}(\kappa_k(\xi') \geq x) = \nu_p(\kappa_k(\xi) \geq x) = \nu_p(\xi \geq x) = \nu_p(\uparrow x).$$

This proves that ν_p is indeed the law of ξ' , and completes the proof of Point 3.

Finally, it remains only to show the following point:

4. The sequence $(Y'_k)_{k \geq 1}$ converges toward ν_p in distribution.

With the same notations as above, we have for every $x \in \mathcal{M}$ and for every integer $k \geq \tau(x)$:

$$\mathbb{P}(Y'_k \geq x) = \mathbb{P}\left(\bigvee_{j \geq 1} Y'_j \geq x\right) = \mathbb{P}(\xi' \geq x) = \nu_p(\uparrow x), \quad \text{by Point 3.}$$

Hence for every $x \in \mathcal{M}$, the value of $\mathbb{P}(Y'_k \geq x)$ is eventually constant when $k \rightarrow \infty$, equal to $\nu_p(\uparrow x)$. This implies the convergence of $(Y'_k)_{k \geq 1}$ in distribution toward the distribution ν_p .

The proof of Theorem 5 is complete.

Proof of Lemma 2. We first show that g is B -invariant on the right. For $p = p_0$, it follows from [2, Prop. 4.12] that the formula $h(c) = p_0^{|c|} g(c)$ holds for all cliques $c \in \mathfrak{C}$. Therefore, for all $c \in \mathfrak{C}$, we have:

$$(Bg)_c = \sum_{c' \in \mathfrak{C} : c \rightarrow c'} p_0^{|c'|} g(c') = \sum_{c' \in \mathfrak{C} : c \rightarrow c'} h(c') = g(c).$$

We now prove that B has spectral radius 1. Let $\|M\|$ denote the spectral radius of a non-negative matrix, that is to say, the largest modulus of its eigenvalues. For $p \leq p_0$, let B_p be the matrix of size $|\mathfrak{C}|$ and defined by

$$(B_p)_{(c,c')} = \begin{cases} 0, & \text{if } \neg(c \rightarrow c') \\ p^{|c'|}, & \text{if } c \rightarrow c'. \end{cases}$$

Hence: $B = B_{p_0}$. Let us show that:

$$\forall p < p_0 \quad \|B_p\| \leq 1. \quad (17)$$

Then by upper semi-continuity of the spectral radius, letting $p \rightarrow p_0$, we will deduce $\|B\| \leq 1$. And since we already proved that B has a right-invariant vector, we will obtain the equality $\|B\| = 1$.

To prove (17), it is enough to show the following:

$$\forall \lambda \in (0, 1) \quad I' \left(\sum_{k \geq 0} (\lambda B_p)^k \right) F < \infty, \quad (18)$$

where I and F are the positive vectors of dimension $|\mathfrak{C}|$ defined as follows:

$$\forall c \in \mathfrak{C} \quad I_c = 1, \quad \forall c \in \mathfrak{C} \quad F_c = \lambda p^{|c|},$$

and I' is the transpose of I . Indeed, (18) implies that $\lambda^{-1} \text{Id} - B_p$ is invertible for all $\lambda \in (0, 1)$, and thus, since B_p is non-negative, that its largest eigenvalue cannot be greater than 1.

Fix $\lambda \in (0, 1)$. Then:

$$\forall k \geq 0 \quad I'(\lambda B_p)^k F = \sum_{u \in \mathcal{M} : \tau(u)=k} \lambda^k p^{|u|},$$

and thus:

$$\begin{aligned} I' \left(\sum_{k \geq 0} (\lambda B_p)^k \right) F &= \sum_{k \geq 0} \lambda^k R_p(k), \\ \text{with } R_p(k) &= \sum_{u \in \mathcal{M} : \tau(u)=k} p^{|u|}. \end{aligned} \quad (19)$$

But, for $p < p_0$, we have:

$$\sum_{k \geq 0} R_p(k) = \sum_{u \in \mathcal{M}} p^{|u|} = G(p) < \infty.$$

Therefore $\lim_{k \rightarrow \infty} R_p(k) = 0$. By (19), it follows that (18) holds, which was to be shown. The proof of Lemma 2 is complete.

Proof of Proposition 3. Clearly, $C_{c,c'} = 0 = P_{c,c'}$ if $c \rightarrow c'$ does not hold. For $c, c' \in \mathfrak{C}$ such that $c \rightarrow c'$ holds, we have:

$$C_{c,c'} = p_0^{|c'|} g(c')/g(c) = h(c')/g(c), \quad (20)$$

by the formula $h(\cdot) = p_0^{|\cdot|} g(\cdot)$ recalled above in the proof of Lemma 2. Still for $p = p_0$, we have, according to formula (5) of Theorem 5:

$$P_{c,c'} = h(c')p_0^{|c|}/h(c) = h(c')/g(c) \quad \text{since } h(c) = p_0^{|c|} g(c). \quad (21)$$

Comparing (20) and (21), we obtain that $P = C$ on $\mathfrak{C} \times \mathfrak{C}$. To get that C is stochastic, it remains only to show that $P_{c,1_{\mathcal{M}}} = 0$ for all cliques $c \in \mathfrak{C}$. And indeed, by formula (5), we have:

$$\forall c \in \mathfrak{C} \quad P_{c,1_{\mathcal{M}}} = \frac{p_0^{|c|} h(1_{\mathcal{M}})}{g(c)} = \frac{p_0^{|c|} \mu_{\mathcal{M}}(p_0)}{g(c)} = 0.$$

The proof of Proposition 3 is complete.

Proof of Theorem 6. Let ν denote the uniform measure on the boundary. We compute the expectation of $\bar{\phi}(C_1 \cdot \dots \cdot C_k)$ under ν as follows:

$$\begin{aligned} \mathbb{E}_{\nu} \bar{\phi}(C_1 \cdot \dots \cdot C_k) &= \sum_{x \in \mathcal{M} : \tau(x)=k} \nu(C_1 \cdot \dots \cdot C_k = x) \left(\sum_{y \in \mathcal{M} : y \leq x \wedge |y|=k} \phi(y) \right) \\ &= \sum_{y \in \mathcal{M} : |y|=k} \phi(y) \left(\sum_{x \in \mathcal{M} : \tau(x)=k \wedge x \geq y} \nu(C_1 \cdot \dots \cdot C_k = x) \right) \\ &= \sum_{y \in \mathcal{M} : |y|=k} \phi(y) \nu(\uparrow y) \\ &= p_0^k \cdot (\#\mathcal{M}_k) \cdot \mathbb{E}_{\nu_{\mathcal{M}_k}} \phi \quad \text{since } \nu(\uparrow y) = p_0^{|y|} = p_0^k \end{aligned}$$

This completes the proof of Theorem 6.